Data Assimilation in Multiscale Systems

Predictability with Stochastic PDEs

Humberto C. Godinez, Balu Nadiga, and J. David Moulton
Los Alamos National Laboratory
Los Alamos, NM 87544

Abstract

Predictive modeling of complex physical systems, such as the climate system, must capture the coupling of phenomena over an ever increasing range of spatial and temporal scales. With an increasing amount of data available for these systems, data assimilation is poised to play a significant role in this modeling. Data assimilation refers to methodologies that use information from the model, observational data, and corresponding error statistics to produce an improved model state, referred to as the analysis.

In this work we explore how the strength of the coupling between scales influences the dynamics of the system, the effectiveness of data assimilation, and the data requirements. We use a simplified model with two different time scales developed by Lorenz (1996). As with the real system, predictability in this model derives from the slow large scales of the system while the faster smaller scales act to reduce predictability. When considering a variation in the coupling between dynamics at the two time scales over a wider range of values than previously considered, we find that the resultant behavior is reminiscent of the main difference in coupling across scales in the atmosphere and ocean systems: The coupling between the large, slow scales (the scale of weather systems; synoptic scales) and the fast convective scales in the atmosphere is relatively weak, where as the coupling between the slow basin scales and the faster mesoscales (roughly the scale of ocean eddies as seen, e.g., in satellite imaging) is strong. When the coupling is strong, the slow scale variable displays variability on the faster scale, qualitatively similar to time series of leading large-scale modes of variability.

We next consider issues of ensemble data assimilation using the two-scale Lorenz system as an example. To mimic existing observational networks for real climate systems, only the slow and large scale variables are observed. To successfully perform data assimilation, in this multiscale context, it is necessary to understand data requirements as a function of coupling strength across scales. In the two-scale Lorenz system, the data-requirements are much more stringent when the coupling across scales is strong. Given a data frequency and sampling that results in an accurate analysis from the data assimilation when the system coupling is weak to moderate, the corresponding analysis is of poor accuracy when the coupling is strong.

Finally we present a new matrix-free implementation of the ensemble Kalman filter (EnKF), which is used in this study. The EnKF is an attractive data assimilation methodology because it evolves the model covariance matrix for a non-linear model, through the use of an ensemble of model states, and avoids the closure problem associated with the traditional extended Kalman filter. However, as more data becomes available, or is necessary for an accurate analysis, a potential bottleneck in the EnKF arises in the matrix inversion implicit in the Kalman gain matrix. In this work we exploit the particular form of the factors of the this matrix, to develop an efficient matrix-free approach. Numerical experiments are presented that demonstrate a significant gain in efficiency over the existing SVD based implementation.